A methodology to avoid over-diversification of funds of equity funds An implementation case study for equity funds of funds in bull markets

Nadège Ribau-Peltre, Pascal Damel, An Lethi

A methodology to avoid over-diversification of funds of equity funds

An implementation case study for equity funds of funds in bull markets

Nadege Peltre CEREFIGE
Pascal Damel LGIPM
An LeThi LGIPM

Abstract

Funds of funds are funds that invest primarily in units of other funds. They have developed in Europe since the end of the 1990s. They exist because no fund manager can be excellent in all fields (all sectors, all geographical areas ...) and because it can therefore be interesting to compose funds from shares of funds managed by different management companies. As there is a higher diversification in funds of funds, they can be attractive at first glance. But studies have pointed out that they have some disadvantages, the main one being over-diversification. In this paper, we will review the literature on the issue of over-diversification by showing the consequences this over-diversification may have on the management and performance of funds of funds. Using Markowitz's mean-variance optimization method, we will on the one hand show that by building funds of funds from a panel of 551 equity funds, the efficient frontier is made up of funds of funds comprising from 1 to 11 funds with an average of 7.44 funds. This empirical study thus shows that the efficient frontier is composed of portfolios comprising a number of funds significantly below the professional standard (20 to 30 funds). Over-diversification and the accumulation of the resulting costs are therefore not a necessity.

On the other hand, we will show that the mean-variance optimization method can be improved by DCA clustering techniques. A prior clustering of the initial database makes it indeed possible to reduce (by almost 10 in our example) the size of the database on which the Markowitz's mean-variance optimization is applied. The efficient frontier deriving from this reduced database is almost equivalent in terms of risk-adjusted performance as the one deriving from the initial database, while avoiding computational problems generated during the optimization process on wide databases (especially when including regulatory constraints).

Key words: funds of funds, Mean-Variance optimization, market expectations, over-diversification, clustering method

1 Introduction

Funds of mutual funds are often considered, in the financial literature, as being part of the family of the mutual funds: Klapper, Sulla et Vittas (Klapper et al., 2004) state that "there are five main types of mutual funds: equity and bond funds that predominantly invest in equities or bonds; balanced funds that have more balanced portfolios of both equities and bonds; money market mutual funds that specialize in short-term instruments; and finally funds of funds that mainly invest in other mutual funds."

The reasons for multi-management are the search for the best managers, a lack of experts or diversification. Multi-management consists in composing an investment fund starting from existing funds. By buying shares of mutual funds, asset managers become in turn customers of other asset managers and diversify their sources of performance in order to reduce risk. This management practice, resulting from alternative management, led to many classical managers being followed. The funds of funds thus strongly developed in Europe at the end of the 1990s, with the emergence of profiled funds, and then a specific regulation framework (in particular with directive UCITS III) was installed.

Fund-of-funds managers sort, analyze and select the best investment funds in the market. Management is based on selectivity and diversification and aims at above-average performance in the reference category for a lower risk-taking.

The existence of funds of funds is based on the principle that, since no fund manager can be excellent in all areas, funds of funds only find their excellence through multi-management, which consists in integrating funds from several management companies (Carretta and Mattarocci, 2005). Since each underlying fund is composed of several securities, the funds of funds are generally highly diversified. This has the effect of significantly reducing the risk of the portfolio.

Funds of funds seem to have everything to seduce investors. However, this management mode has many disadvantages, the common point between them being over-diversification, seen in many funds. Some managers of funds of funds dramatically decrease the profitability of their portfolio by wanting to reduce the risk too much. The sharp decrease in risk through the accumulation of funds in the portfolio can in fact lead to a duplication of funds, the smoothing of the performance and the dilution of profits. In addition, an excessive number of lines is not necessarily optimal when one considers the amount of time spent to manage.

To this is added the additional layer of fees that multi-management requires as a result of the intervention in the process of an additional actor that must be remunerated. The structure of the fund of funds being two floors, it is indeed necessary to pay the companies which manage funds composing the fund of funds, but also to pay the work by the manager of the fund of funds himself.

Based on Markowitz's modern portfolio theory, which is to determine the ideal composition of a portfolio of considered securities, a simplification could be to bring cardinality constraints in the building of a portfolio of funds, in order to avoid over-diversification. However, if the size of the variance-covariance matrix is too large, and, particularly if the number of securities is greater than the number of longitudinal observations of returns, computational problems while optimizing the portfolio composition can occur.

Two alternatives are then possible: increasing the number of temporal observations or reducing the number of assets considered. Increasing the number of observations is often difficult in the building of funds of funds for several reasons. On the one hand, funds have variable lifetimes, and for some "young" funds, it can be impossible to obtain data over a long period. In addition, some studies showed that the variability of the returns of assets being not stationary, integrating a too long period of observations in the study cannot be stated [(MacKinlay, 1987) (Schwert, 1989) (Schwert and Seguin, 1990) (Nawrocki, 1996)]

It is therefore on the second alternative - the reduction of the number of assets in the sample- that we will focus our attention in this paper. In order to decrease the size of the sample, an approach may be to classify the funds and choose the ones of the appropriate classes before any optimization process. In a previous research (Thi et al., 2014), we developped a methodology for classifying funds in three key dimensions in financial portfolios management, namely, risk aversion by investors, investment horizon and anticipation of market trends. We will show that, when selecting only the classes that may be interesting for the fund manager at one point, it is possible to achieve an efficient frontier (according to Markowitz's mean-variance approach) close to the one obtained with a much larger sample.

In our next part, we will draw up a state of the literature on the existence and the consequences of over-diversification of funds of funds, as well as computational problems occurring during the optimization process. In the third part, we will present our research, by indicating our database and the methodology used. The fourth part will be devoted to the presentation and analysis of the results of our study and in the fifth part we will conclude.

2 Literature review

2.1 Literature review on over-diversification

Our literature review will initially focus on the theme of over-diversification of funds, regularly displayed in the financial literature. It is often proved that over-diversification has negative impacts on the returns of funds. This phenomenon, identified for funds, is also found for funds of funds, even if the specific subject of funds of funds - including funds of mutual funds - is less frequently covered in the literature. The consequences of over-diversification on the returns of funds are reflected in three dimensions: the smoothing of returns, the stacking of fees and longer time to manage assets, which limits the quality of managers' analyses. We will conclude this literature review with a summary of scientific works proving the existence and causes of computational problems in mean-variance optimization, especially in a large sample of assets.

Funds of funds have been regularly criticized for several years. Their poor performance is regularly highlighted, both in the public press and in scientific journals. A 2016 study (Peltz International, 2016) based on a survey conducted at the end of 2015 with 32 institutions of hedge funds management, indicates that three-quarters of respondents felt that the situation of these funds had worsened in recent years. However, this finding is not new and Capocci and Nevolo showed in 2005 (Capocci and Nevolo, 2005) that funds of funds tend to underperform their individual components.

Previously, Brown, Goetzmenn and Liang in 2003 (Brown et al., 2003) already reported that individual hedge funds dominate funds of funds on an after-fee return or Sharpe ratio basis. According to Hutson, Lynch and Stevenson (Hutson et al., 2006) funds of funds tend to underperform hedge fund indexes.

Other authors are more moderated and point out that, although funds of funds do not significantly underperform or outperform their benchmark, a trend (certainly statistically insignificant) can be observed for underperformance (Maxam et al., 2007). It is true that this paper relates to funds of hedge funds. But this observation is also true for funds of mutual funds, generally little discussed in the literature. According to a survey of 80 private equity fund-of-funds managers (Prequin Special Report, 2014), when they were asked what the greatest challenges in the funds of funds industry were, the most commonly named issue was returns, selected by 47% of respondents. And the reality is that capital raised for private equity fund-of-funds vehicles is at its lowest levels.

Furthermore, a recent study (Gowri and Deo, 2016) has shown that funds of mutual funds tend to be less efficient than mutual funds. This study attempted to evaluate the performance of funds of funds on the basis of risk-adjusted methods and concluded that that the majority of the equity funds of funds included in the sample had underperformed the benchmark.

Funds of funds therefore appear to have a lower performance than the funds in their category, although a few studies show the opposite ((Gresch and von Wyss, 2011) ; (N et al., 2012) ; (Miller, 2005)). It is quite easy to link this lower performance to the phenomenon of over-diversification regularly highlighted in the literature.

The diversification of a financial portfolio is a basis in market finance, since it allows to undermine financial risks according to the well-known adage "do not put all your eggs in one basket". But too much diversification is not good either, because over-diversification bears its own costs. Holding too many different investments means too much time must be spent tracking them, too much money spent on commissions, and too much mediocrity accepted, diluting the benefit of good investments. According to Sebastian et Attaluri ((Sebastian and Attaluri, 2014), it “has been argued that the relationship between active risk and outperformance (or conversely, over-diversification and underperformance) has developed only since the mid-1990s, and accounts for a large part of the decline in observed skills since that time”.

The phenomenon of over-diversification of funds has already been studied for many years, but studies have focused mainly on mutual funds, hedge funds, funds of hedge funds and funds of private equity funds. For example, Brown et al. in 2011 (Brown et al., 2011) drew attention to a decreasing trend in the returns of funds of hedge funds with the rise in the number of underlying hedge funds in the portfolio. They showed that funds of hedge funds tend to over-diversify their portfolio and be more exposed to left-tail risk. A report of North Sky Capital (North Sky Capital Report, 2011) showed in 2011 that portfolio volatility (standard deviation) decreases as the number of funds and the vintage year diversification are both increased; however, the incremental benefits of diversification decrease with each addition. For example, a portfolio of three underlying funds spread across three vintage years (i.e., one fund in each year) has a standard deviation of 30%, while a 10-fund portfolio and a 20-fund portfolio spread across three vintage years has a standard deviation of 21% and 19%, respectively. This data shows that there is a substantial benefit in investing in a 10-fund portfolio relatively to a 3-fund portfolio, but there is only a modest incremental benefit by investing in a 20-fund portfolio relatively to a 10-fund portfolio. A survey from the research cell of the investment company Cliffwater also shows that a lot of investors tend

to over diversify their hedge fund portfolios, particularly those investing in funds of funds and authors found a negative link between the number of funds in the portfolio and the ability of the fund manager. (Nesbitt and Sokolov, 2015)

We can also include Thomsett (Thomsett, 2014), who, in the preface of his book "Options Trading for the Institutional Investor" indicates that "The three disadvantages mutual funds have over other institutions (insurance companies, banks, pension plans, and specialized advisory service providers, for example) include over-diversification, fees, and cash on hand".

The over-diversification of funds and funds of funds is thus found in mass in the literature. It leads to the smoothing of returns, multiplication of fees and a form of portfolio management made more complex by the increase of assets.

2.1.1 The smoothing of returns

Over-diversification of funds of funds appear like an element leading to the smoothing of their returns. Thus, as noted by Wankhade (Wankhade, 2016): "Although there are many benefits of diversification, there are pitfalls of being over-diversified. Think of it like a sliding scale: the more securities you hold, the less likely you are to feel their individual returns on your overall portfolio. What this means is that though risk will be reduced, so too will the potential for gains." As further noted by Meyer and Mathonet (Meyer and Mathonet, 2011): "An over-diversification will drive returns to a mean, which is in conflict with the main objective of investing in this asset class, namely providing access to top tier funds and generating above average returns". According to Morley (Morley, 2016), if "risk is controlled by diversification – (…) so are returns. (…) Diversification makes theoretical sense, but over-diversification becomes counter-productive." Thus, Greenblat (Greenblat, 2011) reports in his book the words of two fund managers, Angus Gluskie and Atul Lele (managers of Hite Funds Management), who view over-diversification as a common flaw of many fund managers. They also consider that "good investment positions have little impact on overall portfolio returns". Kim et Nelson (Kim and Nelson, 2016) similarly write in 2016 that "when mutual funds add many stocks to their portfolios only to drive down idiosyncratic risk to raise their Sharpe ratios, this can mean that the positive impact of the few great stocks that the fund owns will get watered down."

According to Pomorski (Pomorski, 2009) and Cohen and al. (Cohen et al., 2010) the cause of the poor overall performance of fund managers was not a lack of stock-picking skills, but over-diversification. For Cohen and al., "the poor overall performance of mutual fund managers in the past is not due to a lack of stock-picking ability, but rather to institutional factors that encourage them to over-diversify, i.e. pick more stocks than their best alpha-generating ideas" (Cohen et al., 2010). For Robbé (Robbé, 2008), over-diversification is linked to duplication. And this duplication will lead to the performance of the interests held by the funds of funds in the underlying funds being more strongly correlated to each other and, thus, will undermine the diversification benefits from investing in multiple funds. Abbink (Abbink, 2010) thus considers that "less sophisticated funds of funds managers may simply follow the generally accepted wisdom that exposures should be diversified without considering the ways in which they can hedge each other. This can lead to unintentional hedging away of return potential." Indeed, as Richard Fitzaln Howard noted in 2006 (Jaffer, 2006) when aggregating multiple managers' portfolios, there is the chance that, if mandates overlap, investment positions are doubled or cancelled out. Other authors are more precise in the effects of

over-diversification on the distribution of returns, showing that this one leads to a diminution in excess kurtosis but a slight increase in negative skewness (Mathonet and Meyer, 2008).

2.1.2 The accumulation of fees

There are surprisingly only very few academic studies on the fee structure of funds of funds and the impacts on performance. It should be noted that, as we have previously indicated, the major part of the literature on funds of funds relates to funds of hedge funds. However, as well for hedge funds as for funds of hedge funds, their opacity does not make it possible to collect information on the level of fees and distribution of their costs, in order to generate well-founded scientific analyses. There is however a general agreement that the major disadvantage of the fund-of-funds arrangement is the cost to the investor. In addition to the fees that funds of funds charge, they pass on to the investor all fees charged by the constituent funds. In return for high expenses, funds of hedge funds provide non-performance-based benefits. Some studies however created the controversy of this widespread idea of the low performance of funds of funds: Ang and al. conclude their study published in 2008 with stating that "funds-of-funds, on average, deserve their fees-on-fees"(Ang et al., 2008). But this assessment remains quite scarce and the majority of scientists and experts of the financial industry, agree that "over-diversification—buying more and more mutual funds, index funds, or exchange traded funds—can amplify risk, stunt returns, and increase transaction costs and taxes." (Rekenthaler, 2013). Indeed, "although the total price impact from a single large transaction would be greater than the price impact of multiple smaller transactions, the fixed component of commission fees is higher when multiple trades are made." (Shawky and Smith, 2005). Wu and al (Wu et al., 2017) showed that in real-world applications of portfolio selection models, however, most investors would invest in only a limited number of assets due to market frictions such as management and transaction fees. Brown and al. (Brown et al., 2011) showed that FoFs tend to be over-diversified and that the costly due-diligence from holding so many positions leads to FoFs underperformance. Ang et al. (Ang et al., 2008) already argued that it is the cost of providing due diligence function that justifies the fees that funds of funds charge over the fees charged by constituent funds. Finally, for Debbie Harrison (Harrison, 2005), "over-diversification can lead to unnecessary costs and to style neutrality, at which point it would be simpler and cheaper to use a managed fund or an index tracker".

2.1.3 The difficulty of managing a large portfolio

An over-diversified portfolio is often composed of a large number of assets. As the number of lines increases, difficulties to manage the portfolio increase in the same time. Indeed, over-diversification generates real or fictitious costs that may be similar to opportunity costs. Thus, for Shawky et Smith (Shawky and Smith, 2005) "monitoring costs can take the form of additional personnel costs, as well as poor performance due to a portfolio manager's (or management team's) inability to track a large number of stocks". Ryan (Ryan, 2016) even mentions a "cost of over-diversification". For Ang et al. (Ang et al., 2008) and Brown, Gregoriou, and Pascalau (Brown et al., 2011) the cost of providing a due diligence function particularly justifies the fees that funds of funds charge over the fees charged by constituent funds. Thus, as summarized in Baker and al. (Baker et al., 2015) p 579, "investors should not overemphasize diversification because having too many mutual funds in a portfolio can be costly and time-consuming to monitor".

Thus, if diversification is an important principle in portfolio management, this one should not be thorough to the extreme, at the risk of generating harmful effects on the performance of portfolios. Some well-known authors and managers even claim the benefits of concentration. Note that the funds targeted in this case are primarily hedge funds. As stated by Ridley (Ridley, 2004), "statistical studies suggest that holding many hedge funds provides little benefit by way of reducing portfolio volatility. Investing in too many funds is likely to result in thinner research into each investment. Mediocre research can contribute to increasing risk as opposed to reducing it. A smaller portfolio could be concentrated in higher caliber, better researched funds". According to this author, diversification would be synonymous with mediocrity. This approach has been straightforwardly taken up by the well-known hedge fund manager Warren Buffet, for whom "diversification is what you do when you don't know what to do. (…) Diversification works by spreading risk, but often fails to deliver return. Return only happens when you know what you are doing with a concentrated portfolio" (Morley, 2016). Others, however, are more moderate, and a relatively widespread idea is that a portfolio should be more or less diversified depending on expected market performance.

Indeed, as mentioned by Yeung and al. (Yeung et al., 2012) "concentrated portfolios are geared to performing well when markets are bullish but to performing equally when markets are bearish".

On the assumption that diversification is important, and the observation that over-diversification, on the other hand, has adverse effects on the performance of a portfolio, some authors have begun to investigate whether it is possible to determine the existence of an optimal number of securities in a portfolio.

One of the first studies on the subject was carried out by Evans and Archer in 1968 (Evans John L. and Archer Stephen H., 1968). Those authors showed that a relatively stable and predictable relationship does indeed exist between the number of securities included in a portfolio and the level of portfolio dispersion. Further, this relationship appears to take the form of a rapidly decreasing asymptotic function, with the asymptote approximating the level of systematic variation in the market." The authors also raise doubts concerning the economic justification of increasing portfolio sizes beyond 10 or so securities. In 2013, mutual fund research firm Morningstar (Morningstar.co.uk Editors, 2013) carried out simulations to precisely determine the optimum number of funds in a portfolio. To do this, the researchers computed the standard deviations of portfolios containing 1 to 30 funds over five years. They concluded that holding a single fund is quite risky. Adding a second one substantially reduces volatility, adding another has the same effect. And finally, volatility stabilizes with seven funds. For Brands and Gallagher (Brands and Gallagher, 2005), the majority of diversification benefits are realized when a portfolio of approximately six active equity funds are included in a fund of funds portfolio. John Bogle, the founder of Vanguard (a prominent U.S. fund company), notes in his book Common Sense on Mutual Funds, "I truly believe that it is generally unnecessary to go much beyond four or five equity funds. Too large a number can easily result in over-diversification." (Bogle, 2010). Louis Lowenstein, in his book "The Investor's Dilemma" published in 2008 (Lowenstein, 2008), recommends that mutual fund investors find three, or at most four stock funds that meet their needs.

Finally, Meyer and Mathonet (Mathonet and Meyer, 2008) consider that the most appropriate process for a given manager depends on his/her objectives and tolerance for risks. This means that

trade-offs are inevitable and that a balance must be found between over-diversification and a narrow selection of managers.

The good management of a fund of funds would thus go through the choice of a limited number of funds, which introduces cardinality constraints during the risk-return optimization process, and can generate computational problems, that we will develop in the second paragraph of this literature review.

2.2 Literature review on computational problems occurring during mean-variance optimization process

To create a size-reduced funds of funds when the initial database of funds is large, it is necessary to introduce cardinality constraints to limit the number of assets that compose the portfolio (Lwin et al., 2017). But applying portfolio-theoretic models to a large number of assets is problematic, and the optimization problem is further increased when cardinality constraints are added. Mean-variance portfolio selection consists of finding a self-financing strategy whose final value has maximal mean and minimal variance (Schweizer, 2010). The computational complexity of Markowitz's covariance matrix approach grows exponentially as the number of assets increases, thus increasing computational time and resources (Nawrocki, 1996)

Elton and Gruber (Elton and Gruber, 1973), and DeMiguel et al. (DeMiguel et al., 2009) showed that the estimation of the covariance matrix of the returns of assets is an important step for a successful implementation of the mean-variance portfolio optimization approach; Estimation risks is one of the primary reasons to make standard MV optimization unfeasible in practice (Raimond and Herold, 2006). For Carrasco, and Noumon (Carrasco and Noumon, 2011), two difficulties may occur: the assets could be highly correlated (i.e. the population covariance matrix is nearly singular) or the number of assets could be too large relative to the sample size (i.e. the sample covariance is (nearly) singular even though the population covariance is not). The second difficulty is the one which is the most documented in the literature (see for example (Guo et al., 2017) ; (Kyj et al., 2009); (DeMiguel et al., 2011) ; (Ng, 2014) ; (Jobson and Korkie, 1980); (Ledoit and Wolf, 2003); (Rubio et al., 2012); (Menchero et al., 2011); (Michaud, 2014). Guigues (Guigues, 2011) summarized it as follows: if the number of assets is close to the number of available observations per asset T, then the total number of parameters to estimate is close to the total number of observations, which is problematic. This is due to the phenomenon of non-invertability of the matrix which arises when the number of assets is larger than the number of historical data periods (Senneret et al., 2016). But in addition to non-invertability, the size of the covariance matrix itself could cause error maximization. Having a large number of variance and covariance estimations could potentially lead to estimation outliers resulting in an unbalanced portfolio (Clarke et al., 2006). Some studies even showed that naïve (equally weighted) portfolios can beat mean−variance optimized portfolios (DeMiguel et al., 2009); (Senneret et al., 2016).

To reduce this computational problems, some authors suggested that it was possible to obtain better risk-adjusted performance with fewer securities in the portfolio by using an improved allocation scheme ((Santos, 2015); (Cesarone et al., 2011)).

Yang and Rea (Yang et al., 2015) recommended to reduce the size of an original sample of stocks by applying a PCA (Principal Component Analysis). PCA can be applied to either a covariance matrix or a correlation matrix. The procedure seeked to step-wide remove the most highly

correlated stocks in the sample leaving the most independent stocks. They pointed out that, "from a diversification point of view, eliminating one of a highly correlated pair of stocks results in only a small loss of diversification potential, most of the potential will still be in the sample in the form of the retained stock".

Other authors relied on the clustering statistical technique to facilitate the process of selecting securities and optimizing a fund. In this case, the aim is not to reduce the original database by keeping only the least correlated between the securities, but to classify a large group of securities into a small number of sub-groups made up of relatively homogeneous securities, thereby simplifying the decision making related to these securities. (The Rock Creek Group, Hierarchical Clustering Analysis on the Hedge Funds, December 2010 ; Marvin and Bhatt, 2015). Creating diversified portfolios using cluster analysis, Independent Work Report, Fall 2015, (Souza and Gokcan, 2004)).

This is specifically this method that we also used in the framework of our research, in order to divide our initial (large) sample into sub-samples with identified common financial characteristics (see the first part of this study in: (Thi et al., 2014)).

The literature review of the first paragraph leads us to **our first question**: is it possible to build an efficient fund of funds according to the mean-variance optimization theory, whose number of assets is however limited? And in the continuity of this first question a **second question** emerges: is the performance of a fund of funds correlated to the number of funds included in the portfolio?

Finally, the literature review of the third paragraph leads us to our **third question:** may the efficient curve of funds of funds built through a powerful preliminary ranking of funds and good forecasts of market expectations be equivalent to the efficient curve of funds of funds built on the initial database of funds.

3 Data and methodology

The aim of our study is to show that it is possible to build an optimum efficient frontier with a limited number of funds, in any case smaller than the average number included in the fund of funds in practice, without a significant loss of performance. This helps to limit the accumulation of fees, facilitate the optimization process and integrate the regulatory constraints of any UCITS funds.

3.1. The data base and the clustering technique

Our study was realized on the database constituted of the monthly returns of 551 funds made of European stocks values, in the period from October 2002 to December 2007. For each fund, we calculated, in a first time, 17 performance, risk, and risk-adjusted performance indicators, on a rolling basis of 3 months, 1 year, and 3 years. Our intermediate database was then constituted of 51 variables.

Return measurement	• Absolute performance • Relative performance (to benchmark)
Risk measurement	• Absolute risk (standard deviation) • Tracking Error (relative risk) • Max drawdown • Negative periods (%)

	• Positive periods (%) • VaR Cornish Fisher
Performance measurement adjusted to the risk	• Alphas of Jensen • Betas against the benchmark • Beta bull • Beta bear • Information ratio • Sharpe ratio • Sortino ratio • Treynor ratio • Black Treynor ratio

Table 1. The 17 performance, risk, and risk-adjusted performance indicators used to cluster the funds of our database

In a second time, we computed the four moments (mean, standard deviation, skewness and kurtosis) of each indicator, which witness the 204 variables of our final database.

We have thus applied an ascendant hierarchical classification, with the DCA method to determine the number of classes (Le Thi and al., 2014). We retained 7 classes composed as follows:

Cluster	Number of funds in the cluster	% of funds in the cluster
1	6	1
2	19	3
3	4	1
4	422	77
5	71	13
6	28	5
7	1	0
Sum	551	100

Table 2. The distribution of funds in the 7 generated classes

The financial interpretation of the seven clusters is summarized in the table below:

Classes	Interpretation of the classes
DCA1	The performance of this class is lower than the sample average yield. All funds of this class have fewer negative periods than the average of the sample, but the absolute risk is much higher than the average. With market betas and bull betas above the sample mean, those funds do not behave well on the downside.
DCA 2	The average performance of funds of this class is much higher than that of the sample. This class include funds with positive Jensen alphas, and so an outperformance relative to the market. There were no significant risk indicators for the construction of this class. These funds have a larger variance of the absolute performance and a lower variance for the market beta.
DCA3	The fund's performance of this class is equivalent to that of the sample. This class mainly includes defensive funds, with weaker relative risk and VaR. Nevertheless, we can note that these funds have more negative periods than the sample.
DCA4	With this class, we are in the benchmark sample with a yield which is very close to the average yield of the sample. The risk of all of these funds is slightly lower than that of the sample (Absolute Risk VaR and somewhat lower). We can also note that we did not find any beta or indicators of relative risk among the significant variables.
DCA5	This class includes funds with a lower yield than the average of the sample. But indicators of absolute and relative risks are also lower. As against the VaR is much stronger, which explains the negative return. The risk is poorly estimated by indicators based on the average and standard deviation because returns are clearly nonnormal.
DCA6	The funds in this class have a better than average performance of the sample. This class is just the opposite of the class DCA5, in which the funds have higher aboslute and relative risks, but a lower VaR. Here the non-normality plays for the fund's performance.
DCA7	This class is composed of a single fund. We note that the performance is generally very negative and risk indicators very high

Table 3. The financial description the 7 generated classes

We then applied Markowitz's mean-variance optimization method on 3 datasets: the set of all 551 funds of our initial database, the set of the funds of the DCA2 and DCA4 clusters, behaving well in bullish market (which represents 47 funds), and the set of the funds of the DCA3 and DCA5 clusters, behaving well in bearish market (which represents 75 funds).

Let

nF be the number of assets available (551 in our application);

$\bar{r}_i$ be the expected return of asset i;

σ_{ij} be the covariance between assets i and j;

r* be the desired expected return.

Our decision variables are:

w_i the proportion held of asset i in the portfolio.

The standard Mean-Variance Markowitz model (Markowitz, 1952) finds the proportion vector w that solves the following minimization problem

minimize

$$\sum_{i=1}^{nF} \sum_{j=1}^{nF} \mathrm{w}_i \mathrm{w}_j \sigma_{ij}$$

subjet to

$\sum_{i=1}^{nF} \mathrm{w}_i \bar{\mathrm{r}}_i = r^*$,

$$\sum_{i=1}^{nF} w_i = 1$$

$0 \leq \mathrm{w}_i$, i = 1,…, nF

Note that for a sequence of proportion wi, i=1,...,nF, the risk associated with the portfolio is the corresponding total variance

$$\sum_{i=1}^{nF} \sum_{j=1}^{nF} \mathrm{w}_i \mathrm{w}_j \sigma_{ij}$$

and the return of the portfolio is given by

$$\sum_{i=1}^{nF} \mathrm{w}_i \bar{\mathrm{r}}_i$$

For a fixed value of r*, this problem is a convex quadratic programming problem (QP) for which efficient algorithms exist to compute solutions. By resolving the above QP for varying values of r*, we can trace out the efficient frontiers, a smooth non-decreasing curve that gives the best possible tradeoff of risk against return.

Let us recall some computations:

$P_0^i, P_1^i, \ldots, P_N^i$, N+1 the price of asset I, i=1,…, nF

$$\mathrm{r}_k^i = \frac{\mathrm{P}_k^i - \mathrm{P}_{k-1}^i}{\mathrm{P}_{k-1}^i} \times 100, k = 1, \ldots N$$

$\overline{r^i}$i= $\sum_{k=1}^{N} r_k^i$ i=1,… nF

$\overline{\sigma_{ij}} = (\frac{1}{N-1}) \sum_{k=1}^{N} (\mathrm{r}_k^i - \overline{r^i})\ (\mathrm{r}_k^j - \overline{r^j})$, i=1,…, nF, j=1,…,nF

By varying the desired expected return r* in the interval $[r_{min}^*; r_{max}^*]$ and solving the corresponding optimization problem, we plot the efficient frontier. In our simulations, we have

discretized this interval with the step $\frac{r^*_{max} - r^*_{min}}{1000}$ (because we computed 1000 portfolios in each case)

In the standard Markowitz model, we have $r^*_{max} = \sum_{i=1}^{nF} w_i^{max} \bar{r}_i$ with

$$w^{max} = Argmax \sum_{i=1}^{nF} w_i \bar{r}_i$$

subject to

$$\sum_{i=1}^{nF} w_i = 1$$

$0 \leq w_i$, i = 1,…, nF

and $r^*_{min} = \sum_{i=1}^{nF} w_i^{min} \bar{r}_i$ wi with

$$w^{min} = Argmin \sum_{i=1}^{nF} \sum_{j=1}^{nF} w_i w_j \sigma_{ij}$$

subject to

$$\sum_{i=1}^{nF} w_i = 1$$

$0 \leq w_i$, i = 1,…, nF

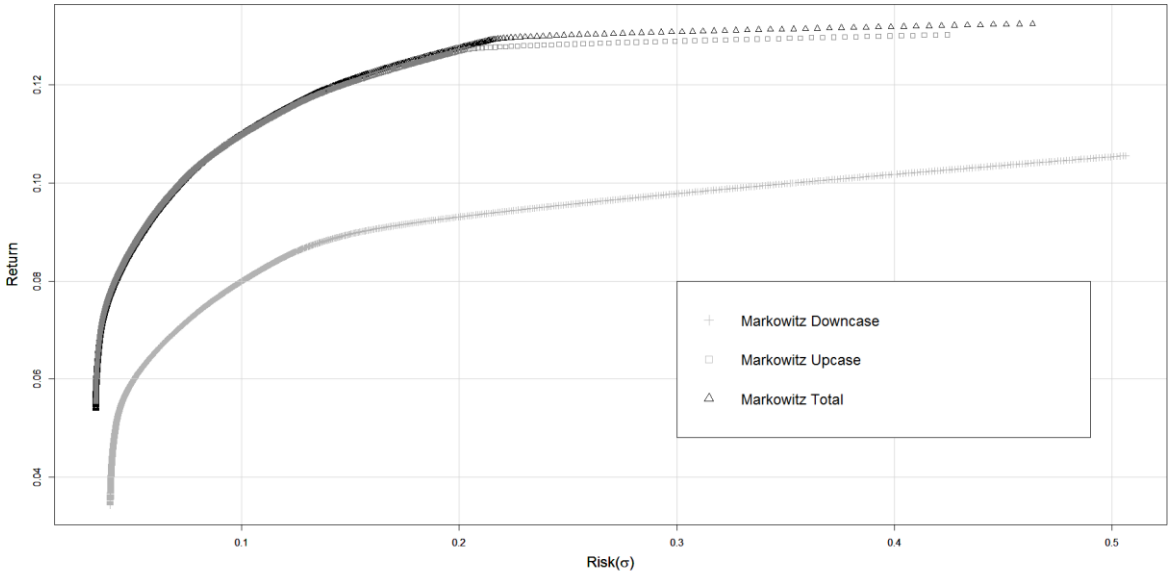
3.2. The three efficient frontiers

We obtained three efficient frontiers, as follows:

Markowitz-Total: corresponds to the standard Markowitz model applied directly on the initial database (551 funds),

Markowitz-Upcase : corresponds to the standard Markowitz model applied on the upcase data produced with the clustering process (the funds of the DCA2 and DCA4 classes which represents 47 funds)

Markowitz-Downcase: corresponds to the standard Markowitz model applied on the downcase data produced with the clustering process (the funds of the DCA3 and DCA5classes, which represents 75 funds).



Graph 1. Efficient frontiers on the series Markowitz-Total, Markowitz-Upcase and Markowitz-Downcase

We can observe that, whatever the set of funds, the number of observations (1368 daily returns between October 2002 and December 2007) is far greater than the number of funds in the optimization process (551 funds for the overall database, 47 funds in the assumption of bull market and 75 funds in the assumption of bear market). Problems of matrix inversion traditionally met in this kind of optimization, when the number of data is higher than the number of observations, do not appear in this study.

4 Results: analyses

4.1 A limited number of funds is enough when creating a fund of funds

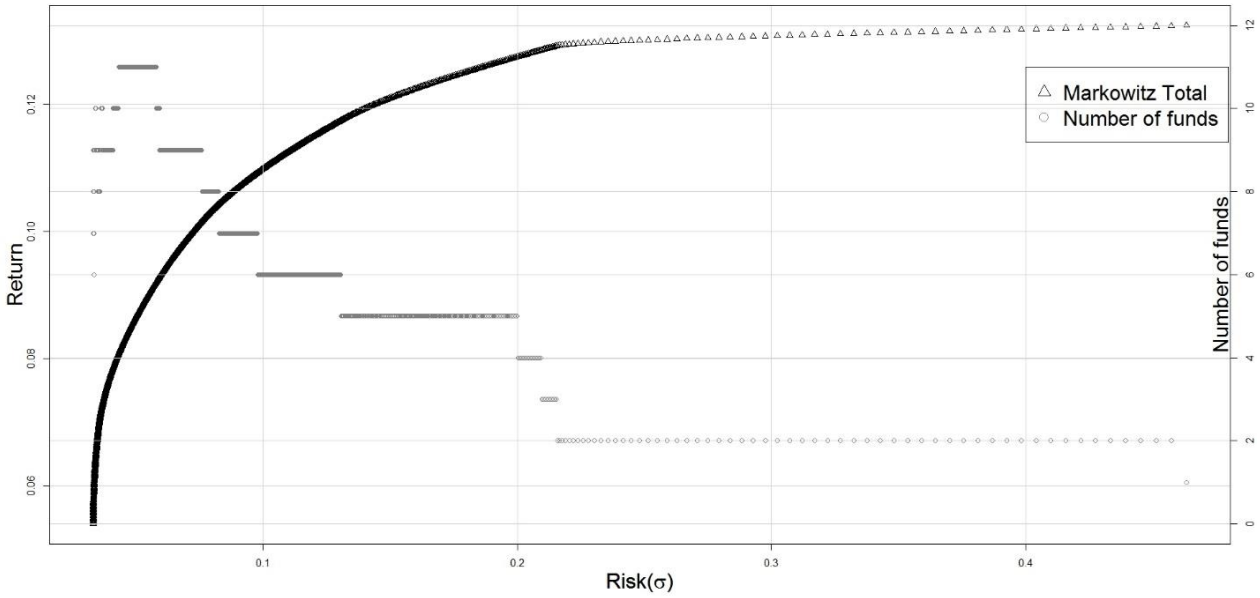
The efficient frontier constituted from the initial database (Markowitz-Total curve) is in reality made up of funds of funds comprising between 1 and 11 funds, with an average of 7.44 funds for Markowitz-Total, 7.42 funds for Markowitz Upcase and 6.37 funds for Markowitz Downcase.

There are in fact the statistical characteristics of the sample of 1000 portfolios constituting the efficient frontier:

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Number.of.funds.Total	1,000	7.437	2.493	1	5	8	9	11
Number.of.funds.Upcase	1,000	7.415	2.495	1	5	8	9	11
Number.of.funds.Downcase	1,000	6.373	2.981	1	3	8	9	11

Table 4. Univariate statistics of the variable Number of funds

We can also visualize how many funds the 1000 portfolios of the efficient frontier are composed of:



Graph 2. Efficient frontier Markowitz-Total returns and number of funds according to risk level

The number of funds in a portfolio does not decrease linearly with the increasing risk: in fact, it increases very quickly in the first (most convex) part of the efficient frontier and then decreases in stages (the same shape of graph can be visualized in Appendix 1 for Markowitz-Upcase and Markowitz-Downcase).

The maximum number of 11 funds in a portfolio on Markowitz-Total is reached for a total of 133 portfolios with an average return of 8,66% and an average volatility of 5.01%. (cf. Table 5).

We have highlighted that the number of funds included in the portfolios of the efficient frontier (7.44 on average, and 11 at most), is well below the number of funds included in the fund of funds, in managers' practice.
Markowitz's mean-variance optimization method can then contribute to help reduce management costs by limiting the number of funds in the funds of funds.
On the other hand, we found that a prior clustering of an initial database of funds, based on good market anticipations, can help reduce the size of the database (that is the space of possible), thus facilitating the mean-variance optimization process.

4.2 An initial sorting of the funds makes it possible to reach a high efficient frontier.

2002 to 2007 is a period of growth for economy and finance and the majority of the financial markets was bullish. It is therefore not surprising that Markowitz-Total and Markowitz-Upcase series are in close proximity to each other. Markowitz Downcase series is graphically clearly below Markowitz-Total and Markowitz-Upcase series.

Since Markowitz-Total and Markowitz-Upcase curves are very close to each other, it is visually difficult to determine if one is above the other. Since abscissas of the two curves are different, we used interpolation methods to determine the extent to which the two curves differ from each other.

We used two interpolation methods: the linear interpolation method (see Appendix 6) and the spline cubic method (see appendix 7). We interpolated Markowitz-Total series with these two methods, and compared the interpolated series with the raw data of Markowitz-Upcase series by calculating the point-by-point differences and the average difference between points of the two series.

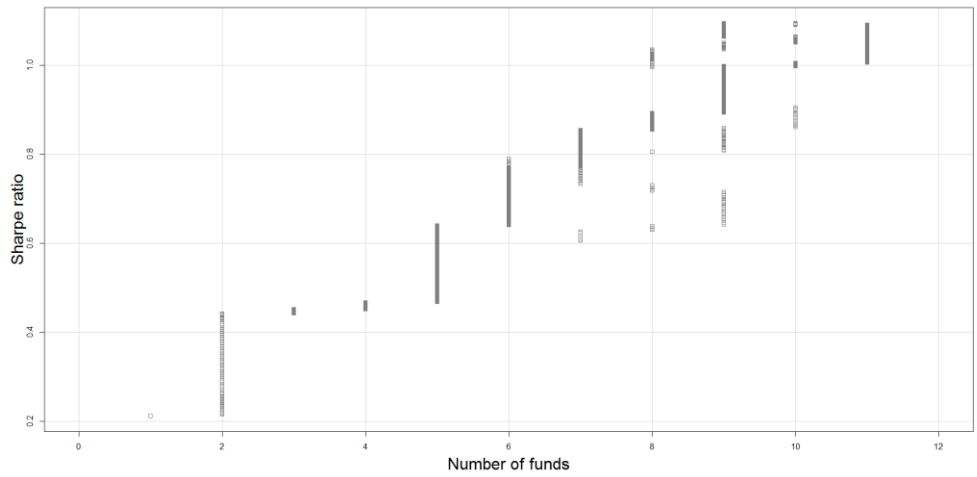
The point-by-point differences are positive 46.1 % of the time and negative 53.9 % of the time for both methods. The average of the differences are positive using the two methods: the average difference is $6{,}147.10^{-5}$ for the first method, and $6.153.10^{-5}$ % for the second.

These results show that Markowitz-Downcase curve is below and far from Markowitz-Toral curve, and that Markowitz-Upcase curve appears very close to Markowitz-Total curve (sometimes below this latter, sometimes above, even if the average difference between the two series is positive, but very weak).

4.3 Sharpe ratios of the optimal fund of funds can be modeled according to the number of funds included in the portfolio

We showed that the number of funds included in the efficient frontier was much lower than the number of funds in the initial database, and than the number of funds generally included in the practice (from 20 to 30 funds). We wanted to take the analysis further, asking whether there was a correlation between the risk-adjusted level of performance of efficient frontier portfolios and the number of funds included in them. We then computed the Sharpe ratio for each of the 1000 portfolios for the 3 series Markowitz-Total, Markowitz-Upcase and Markowitz-Downcase.

For the risk-free rate, we used the average daily performance of 12 months-EURIBOR, computed from the average annual rate over the period under review (October 2002-December2007).



Graph 3. Sharpe ratio of the 1000 portfolios of Markowitz-Total according to their number of funds

We can notice that Sharpe ratio tends to increase (non-linearly however) with the number of funds for Markowitz-Total (the same trend can be visualized in Appendix 2 for Markowitz-Upcase and

Markowiz-Downcase). The maximum Sharpe ratio (1.053) for Markowitz-Total is achieved for a portfolio of 11 funds (with a mean return of 8.66 % and a risk of 5.01 %).

The following table summarizes statistical on the 1000 portfolios according to the number of funds in the portfolio, for Markowitz-Total (the same table can be visualized in Appendix 3 for Markowitz-Upcase and Markowiz-Downcase). :

Number of funds in the portfolio	Number of portfolios	Risk (%)	Return (%)	Sharpe ratio
1	1	46,35%	13,25%	0,213
2	57	31,37%	13,09%	0,326
3	12	21,23%	12,89%	0,447
4	18	20,46%	12,81%	0,460
5	168	16,22%	12,26%	0,553
6	132	11,05%	11,14%	0,706
7	94	7,65%	9,54%	0,792
8	67	5,98%	8,71%	0,898
9	255	5,01%	8,17%	0,953
10	63	4,15%	7,64%	1,019
11	133	5,01%	8,66%	1,053
Overall	**1000**	**9,97%**	**9,87%**	**0,800**

Table 5. Statistical of the 1000 portfolios according to the number of funds in the portfolio for Markowitz-Total

We can notice that if the risk and return of the portfolios decrease with the number of funds in the portfolio, the Sharpe ratio tends to grow in the same time.

Let us now turn to the correlation between the four variables mentioned above: return, risk, Sharpe ratio and number of funds, for Markoowitz-Total (the correlations for Markowitz-Upcase and Markowitz-Downcase can be visualized in Appendix 4):

	Markowitz.Total.Risk	*Markowitz.Total.Return*	*Sharpe.ratio*	*Number.of.funds*
Markowitz.Total.Risk				
Markowitz.Total.Return	0.821***			
Sharpe.ratio	-0.876***	-0.732***		
Number.of.funds	-0.863***	-0.769***	0.935***	

Computed correlation used pearson-method with listwise-deletion.

*Note: * = p < .05, ** = p < .01, *** = p < .001*

Table 6. Matrix correlation on the variables relative to Markowitz-Total

Correlation is strong and positive between return and risk, and relatively strong but negative between return and Sharpe ratio, which may seem surprising at first. However, the Sharpe ratio of

any security is composed of both its return (which appears in the numerator) and its risk (in the denominator). And on the right-hand side of the efficient frontier, risk increases much more rapidly than return. It explains the global negative correlation between Sharpe ratio and return, with the Sharpe ratio declining, even as return increases, because risk increases much more rapidly than return does.

Considering the number of funds, it is strongly, but negatively correlated with return, which means that return increases when the number of funds decreases. On the other hand, the number of funds is high, but positively correlated with the Sharpe ratio, which means that the Sharpe ratio tends to become larger as the number of funds increases.

It emerges from these observations that the risk-adjusted performance of a portfolio of funds depends on the number of funds included in the portfolio. We therefore modeled the Sharpe ratio as a function of the number of funds in the portfolio, using a simple linear regression.

We tested the effects of the number of funds on Sharpe ratio level using the following linear regression model, for the three efficient frontiers (Markowitz-Total, Markowitz-Upcase and Markowitz-Downcase).

We used two specification models, one with only one explicative variable, the number of funds for the portfolios of the efficient frontier, and the other including two control variables, risk and return for the portfolios of the efficient frontier (the univariate statistic of the variables included in the regressions is on Appendix 5).

First model is specified as follow for each portfolio i of an efficient frontier:

$$Sharpe.ratio_i = c + \alpha.\ Number.of.funds_i + \epsilon_i$$

Second model is specified as follow for each portfolio i of an efficient frontier:

$$Sharpe.ratio_i = c + \alpha.\ Number.of.funds_i + Risk_i + Return_i + \epsilon_i$$

Results are the following:

	Dependent variable:					
	Sharpe.ratio.Total		Sharpe.ratio.Upcase		Sharpe.ratio.Downcase	
	(1)	(2)	(3)	(4)	(5)	(6)
Number.of.funds.Total	0.085***	0.066***				
	(0.001)	(0.002)				
Markowitz.Total.Return		0.993***				
		(0.189)				
Markowitz.Total.Risk		-0.996***				
		(0.070)				
Number.of.funds.Upcase			0.085***	0.064***		
			(0.001)	(0.002)		

Markowitz.Upcase.Return				0.648***		
				(0.177)		
Markowitz.Upcase.Risk				-1.071***		
				(0.069)		
Number.of.funds.Downcase					0.030***	0.053***
					(0.001)	(0.001)
Markowitz.Downcase.Return						9.746***
						(0.152)
Markowitz.Downcase.Risk						-0.855***
						(0.025)
Constant	0.166***	0.310***	0.185***	0.383***	0.170***	-0.591***
	(0.008)	(0.027)	(0.007)	(0.023)	(0.008)	(0.017)
Observations	1,000	1,000	1,000	1,000	1,000	1,000
R^2	0.875	0.896	0.894	0.917	0.409	0.889
Adjusted R^2	0.874	0.896	0.894	0.916	0.408	0.889
Residual Std. Error	0.081 (df = 998)	0.073 (df = 996)	0.073 (df = 998)	0.065 (df = 996)	0.107 (df = 998)	0.046 (df = 996)
F Statistic	6,954.521*** (df = 1; 998)	2,865.204*** (df = 3; 996)	8,402.321*** (df = 1; 998)	3,653.198*** (df = 3; 996)	689.744*** (df = 1; 998)	2,654.576*** (df = 3; 996)
Note:						*p<0.1; **p<0.05; ***p<0.01

We can note that the R^2coefficient of determination in the first model is high for Markowitz-Total and Markowitz-Upcase, but much lower for Markowitz-Downcase. Introducing control variables risk and return leads to a small increase of R^2 for Markowitz-Total and Markowitz-Upcase, but a sharp increase of Markowitz-Downcase.

These results seem to mean that for Markowitz-Total and Markowitz Upcase, the Sharpe ratio depends on the number of funds included in the portfolio, but this seems not to be really the case for Markowitz-Downcase. Sharpe ratio increases in average of 0.066 when a fund is added in the portfolio for Markowitz-Total, and of 0.064 for Markowitz-Upcase, which is a very close value.

5 Conclusion

The results show that the two Markowitz-Total and Markowitz-Upcase curves are very close to each other, the Markowitz-Total series is only very slightly higher than the Markowitz-Upcase series, which means that, for the same level of risk, only a very slightly higher level of return is obtained with a portfolio composed from the 551 equity funds in the original database, than with a portfolio made up of 47 funds that behave well in bull market. On the other hand, the Markowitz-Downcase curve is clearly well below the Markowitz-Total curve, which shows that the funds used to build this efficient frontier are not adequate in a bear market context.

We can thus conclude that it is possible to reach almost the same efficient frontier from a sample of 47 funds than from a sample of 551 funds, if the funds have been correctly selected to be in line

with market expectations. It is therefore possible to build funds in an optimal way by reducing management costs.

In addition, our results are in line with studies related to the ideal number of securities in a fund. We indeed showed that the portfolios of the efficient frontiers Markowitz-Total and Markowitz-Upcase are composed of between 2 and 11 funds, with an average of 7.44 funds for Markowitz-Total and 7.42 funds for Markowitz-Upcase. We also showed that for these two efficient frontiers, the Sharpe ratio tends to increase with the number of funds included in the portfolio. We might want to conclude from our results that the ideal number of funds of funds is therefore 11 funds, much fewer than the professional average (20 to 30 funds).

However, it is really time to recall now that the results of this research have operational limits. The first is linked to the possibility to anticipate the possibilities of market expectations. Our method is indeed based on market expectations, which determine the selection of funds, and it is on this selection that Markowitz's method of average-variance optimization is applied. However, since the 2007-2008 crisis, as markets have been particularly disrupted, it has become even more difficult than before to create market expectations, even for financial specialists. The risk is that, if expectations prove to be wrong, one could make inaccurate choices and include in the selection, funds that could not in any case generate a high efficient frontier. The second is linked to the size of the sampling: we realized this study on a limited number of funds (551 funds in our database). This certainly allowed us to avoid the computational problems related to large samples (see paragraph 2.2) but it would be interesting to carry out this study on several samples of this size, in order to verify the reproducibility of our results.

References

Abbink, J.B. (2010). Alternative Assets and Strategic Allocation: Rethinking the Institutional Approach (John Wiley & Sons).

Ang, A., Rhodes-Kropf, M., and Zhao, R. (2008). Do Funds-of-Funds Deserve Their Fees-on-Fees? (Cambridge, MA: National Bureau of Economic Research).

Baker, H.K., Filbeck, G., and Kiymaz, H. (2015). Mutual Funds and Exchange-traded Funds: Building Blocks to Wealth (Oxford University Press).

Bogle, J.C. (2010). Common Sense on Mutual Funds (John Wiley & Sons).

Brands, S., and Gallagher, D.R. (2005). Portfolio selection, diversification and fund-of-funds: a note. Account. Finance *45*, 185–197.

Brown, S.J., Goetzmann, W.N., and Liang, B. (2003). Fees on Fees in Funds of Funds (National Bureau of Economic Research).

Brown, S.J., Gregoriou, G.N., and Pascalau, R.C. (2011). Diversification in Funds of Hedge Funds: Is it Possible to Overdiversify? (Rochester, NY: Social Science Research Network).

Capocci, D.P.J., and Nevolo, V. (2005). Funds of Hedge Funds versus Portfolios of Hedge Funds - A Comparative Analysis (Rochester, NY: Social Science Research Network).

Carrasco, M., and Noumon, N. (2011). Optimal portfolio selection using regularization (Discussion paper).

Carretta, A., and Mattarocci, G. (2005). Funds of Funds' Portfolio Composition and its Impact on Performance: Evidence from the Italian Market.

Cesarone, F., Scozzari, A., and Tardella, F. (2011). Portfolio selection problems in practice: a comparison between linear and quadratic optimization models. ArXiv11053594 Q-Fin.

Cohen, R.B., Polk, C., and Silli, B. (2010). Best Ideas. SSRN Electron. J.

DeMiguel, V., Garlappi, L., Nogales, F.J., and Uppal, R. (2009). A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. Manag. Sci. *55*, 798–812.

DeMiguel, V., Martin-Utrera, A., and Nogales, F.J. (2011). Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection (Rochester, NY: Social Science Research Network).

Elton, E.J., and Gruber, M.J. (1973). Estimating the Dependence Structure of Share Prices — Implications for Portfolio Selection. J. Finance *28*, 1203–1232.

Evans John L., and Archer Stephen H. (1968). Diversification and the reduction of dispersion: an empirical analysis. J. Finance *23*, 761–767.

Gowri, M., and Deo, M. (2016). Performance Evaluation of Equity Oriented Growth and Dividend Funds of Mutual Funds in India : An Application of Risk – Adjusted Theoretical Parameters. Indian J. Finance *10*, 43–54.

Greenblat, E. (2011). Young Guns on the Sharemarket: Meet Australia's New Generation of Money Makers (John Wiley & Sons).

Gresch, N., and von Wyss, R. (2011). Private Equity Fund of Funds vs. Funds : A Performance Comparison. J. Priv. Equity *14*, 43–58.

Guigues, V. (2011). Sensitivity analysis and calibration of the covariance matrix for stable portfolio selection. Comput. Optim. Appl. *48*, 553–579.

Guo, S., Box, J.L., and Zhang, W. (2017). A Dynamic Structure for High-Dimensional Covariance Matrices and Its Application in Portfolio Allocation. J. Am. Stat. Assoc. *112*, 235–253.

Harrison, D. (2005). Personal Financial Planning: Theory and Practice (Pearson Education).

Hutson, E., Lynch, M., and Stevenson, M. (2006). 14 - Distributional characteristics of funds of hedge funds returns. In Funds of Hedge Funds, G.N. Gregoriou, ed. (Boston: Butterworth-Heinemann), pp. 213–238.

Jaffer, S. (2006). Multi-manager Funds: Long-only Strategies for Managers and Investors (Euromoney Books).

Jobson, J.D., and Korkie, B. (1980). Estimation for Markowitz Efficient Portfolios. J. Am. Stat. Assoc. *75*, 544–554.

Kim, K., and Nelson, W. (2016). Mutual funds exposed (EQIS Capital Management, Inc).

Klapper, L., Sulla, V., and Vittas, D. (2004). The development of mutual funds around the world. Emerg. Mark. Rev. *5*, 1–38.

Kyj, L., Ostdiek, B., and Ensor, K. (2009). Realized covariance estimation in dynamic portfolio optimization (Technical Report, Working Paper, Humboldt-Universit).

Ledoit, O., and Wolf, M. (2003). Honey, I Shrunk the Sample Covariance Matrix (Rochester, NY: Social Science Research Network).

Lowenstein, L. (2008). The Investor's Dilemma: How Mutual Funds Are Betraying Your Trust And What To Do About It (John Wiley & Sons).

Lwin, K.T., Qu, R., and MacCarthy, B.L. (2017). Mean-VaR portfolio optimization: A nonparametric approach. Eur. J. Oper. Res. *260*, 751–766.

MacKinlay, A.C. (1987). On multivariate tests of the CAPM. J. Financ. Econ. *18*, 341–371.

Markowitz, H. (1952). Portfolio selection. J. Finance *7*, 77–91.

Marvin, K., and Bhatt, S. Creating Diversified Portfolios Using Cluster Analysis - Independent Work Report Fall, 2015,
https://www.cs.princeton.edu/sites/default/files/uploads/karina_marvin.pdf

Mathonet, P.-Y., and Meyer, T. (2008). J-Curve Exposure: Managing a Portfolio of Venture Capital and Private Equity Funds (John Wiley & Sons).

Maxam, C.L., Ong, S.E., and Wisen, C. (2007). Funds of Funds: Diversification, Selection or Expense Arbitrage? In Diversification and Portfolio Management of Mutual Funds, G.N. Gregoriou, ed. (London: Palgrave Macmillan UK), pp. 18–56.

Menchero, J., Wang, J., and Orr, D.J. (2011). Eigen-Adjusted Covariance Matrices (Rochester, NY: Social Science Research Network).

Meyer, T., and Mathonet, P.-Y. (2011). Beyond the J Curve: Managing a Portfolio of Venture Capital and Private Equity Funds (John Wiley & Sons).

Michaud, R.O. (2014). The Markowitz Optimization Enigma: is "Optimized" Optimal? (Rochester, NY: Social Science Research Network).

Miller, D.E. (2005). The Risk Profile of Private Equity Funds of Funds. CFA Dig. *35*, 13–14.

Morley, I. (2016). Morley's Laws of Business and Fund Management (Paragon Publishing).

N, V., S, M., and K, C.S.R. (2012). The Relationship Between Fund Performance and Fund Characteristics: Evidence from India. IUP J. Appl. Finance *18*, 5–18.

Nawrocki, D. (1996). Portfolio analysis with a large universe of assets. Appl. Econ. *28*, 1191–1198.

Nesbitt, S.L., and Sokolov, El.J. (2015). How Many Hedge Funds? Only 15-20 For The Highly Skilled | Hedge Fund Insight.

Ng, K.W. (2014). Can random matrix theory resolve Markowitz optimization enigma? (No Publisher Supplied).

North Sky Capital Report (2011). North sky capital - alternative assets and the use of fund of funds.

Peltz international "hedge funds" 2016 pdf worsened - Recherche Google.

Pomorski, L. (2009). Acting on the Most Valuable Information: "Best Idea" Trades of Mutual Fund Managers (Rochester, NY: Social Science Research Network).

Raimond, M., and Herold, U. (2006). Portfolio Choice and Estimation Risk. A Comparison of Bayesian to Heuristic Approaches. ASTIN Bull. *36*, 135–160.

Rekenthaler, J. (2013). Can You Be Too Diversified? | Morningstar.

Ridley, M. (2004). How to Invest in Hedge Funds: An Investment Professional's Guide (Kogan Page Publishers).

Robbé, J.J. de V. (2008). Securitization Law and Practice: In the Face of the Credit Crunch (Kluwer Law International).

The Rock Creek Group, Hierarchical Clustering Analysis on the Hedge Funds, December 2010

Rubio, F., Mestre, X., and Palomar, D.P. (2012). Performance Analysis and Optimal Selection of Large Minimum Variance Portfolios Under Estimation Risk. IEEE J. Sel. Top. Signal Process. *6*, 337–350.

Ryan, F.X. (2016). Revolutionizing Accounting for Decision Making: Combining the Disciplines of Lean with Activity Based Costing (Xlibris).

Santos, A.A.P. (2015). Beating the market with small portfolios: Evidence from Brazil. EconomiA *16*, 22–31.

Schweizer, M. (2010). Mean-Variance Hedging. In Encyclopedia of Quantitative Finance, R. Cont, ed. (Chichester, UK: John Wiley & Sons, Ltd), p.

Schwert, G.W. (1989). Why Does Stock Market Volatility Change Over Time? J. Finance *44*, 1115–1153.

Schwert, G.W., and Seguin, P.J. (1990). Heteroskedasticity in Stock Returns. J. Finance *45*, 1129–1155.

Sebastian, M., and Attaluri, S. (2014). Conviction in Equity Investing.

Senneret, M., Malevergne, Y., Abry, P., Perrin, G., and Jaffres, L. (2016). Covariance Versus Precision Matrix Estimation for Efficient Asset Allocation. IEEE J. Sel. Top. Signal Process. *10*, 982–993.

Shawky, H.A., and Smith, D.M. (2005). Optimal Number of Stock Holdings in Mutual Fund Portfolios Based on Market Performance. Financ. Rev. *40*, 481–495.

Souza, C.D., and Gokcan, S. (2004). Allocation Methodologies and Customizing Hedge Fund Multi-Manager Multi-Strategy Products. J. Altern. Invest. *6*, 7–22.

Thi, H.A.L., Damel, P., Peltre, N., and Phuc, N.T. (2014). The Confrontation of Two Clustering Methods in Portfolio Management: Ward's Method Versus DCA Method. In Advanced Computational Methods for Knowledge Engineering, (Springer, Cham), pp. 87–98.

Thomsett, M.C. (2014). Options trading for the institutional investor: managing risk in financial institutions (Upper Saddle River, NJ: Pearson Education).

Wankhade, C. (2016). Mutual Funds For All: Mutual Funds For All (Mutual Funds For All).

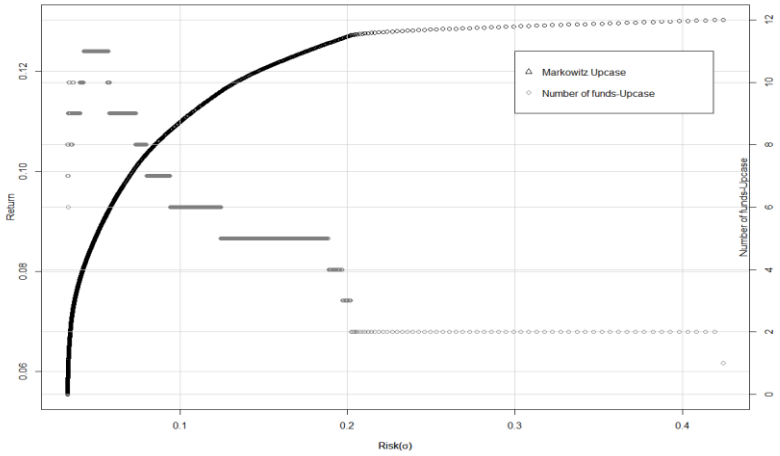
Wu, B., Sun, X., Li, D., and Zheng, X. (2017). Quadratic Convex Reformulations for Semicontinuous Quadratic Programming. SIAM J. Optim. *27*, 1531–1553.

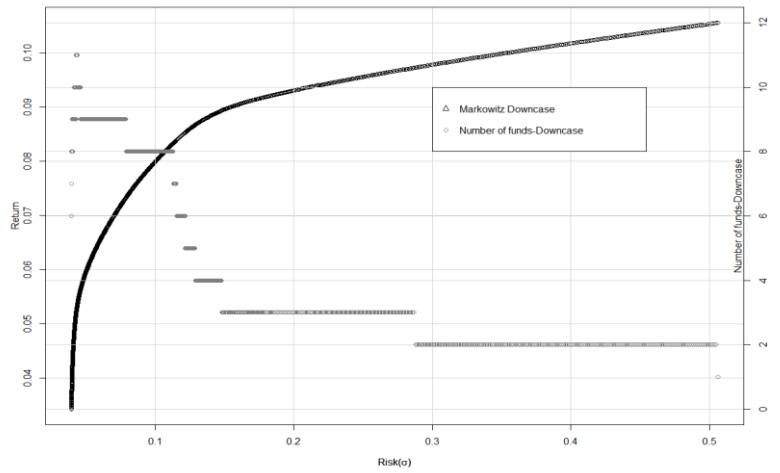
Yang, L., Rea, W., and Rea, A. (2015). Stock Selection with Principal Component Analysis (University of Canterbury, Department of Economics and Finance).

Yeung, D., Pellizzari, P., Bird, R., Abidin, S., and others (2012). Diversification versus Concentration... and the Winner is?

Appendix 1

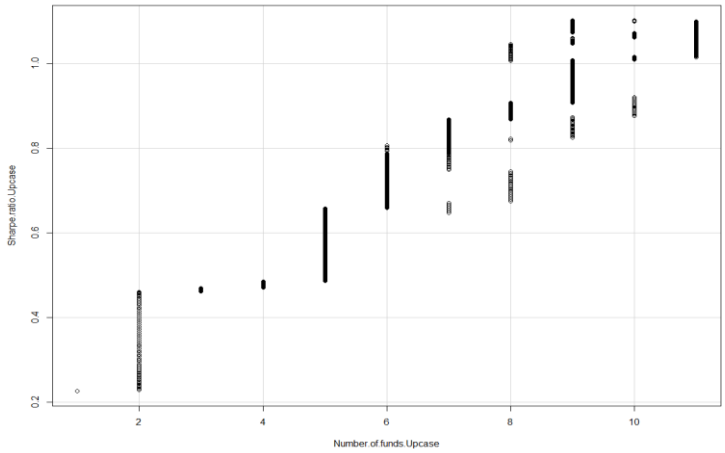
Evolution of the returns and the number of funds for Markowitz-Upcase and Markowitz-Downcase efficient frontiers, according to the risk level

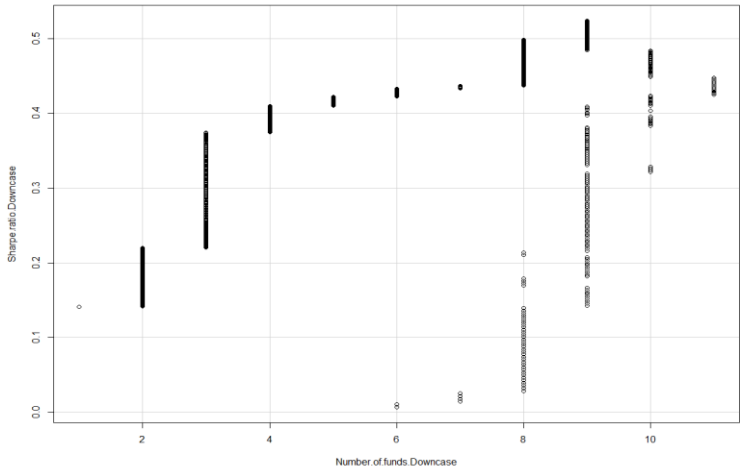




Appendix 2

Graphs of the Sharpe ratios of portfolios of Markowitz-Upcase and Markowitz-Downcase, according to the number of funds included in the portfolio





Appendix 3

Statistical on the 1000 portfolios according to the number of funds in the portfolio, for Markowitz-Upcase and Markowitz -Downcase

For Markowitz-Upcase :

Number of funds in the portfolio	Number of portfolios	Risk (%)	Return (%)	Sharpe ratio
1	1	42,43%	13,03%	0,227
2	58	28,96%	12,87%	0,344
3	11	19,95%	12,69%	0,466
4	18	19,27%	12,61%	0,478
5	170	15,38%	12,09%	0,571
6	132	10,61%	11,02%	0,724
7	95	7,39%	9,44%	0,807
8	83	5,42%	8,19%	0,881
9	234	5,08%	8,38%	0,988
10	62	4,09%	7,61%	1,028
11	136	4,94%	8,64%	1,063
Overall	**1000**	**9,57%**	**9,82%**	**0,816**

For Markowitz-Downcase :

Number of funds in the portfolio	Number of portfolios	Risk (%)	Return (%)	Sharpe ratio
1	1	50,63%	10,56%	0,141
2	168	39,39%	10,15%	0,175
3	148	20,98%	9,34%	0,292
4	47	13,77%	8,80%	0,393
5	24	12,50%	8,60%	0,416
6	26	11,23%	8,07%	0,396
7	14	9,28%	6,96%	0,316
8	197	8,39%	7,02%	0,394
9	311	5,50%	5,95%	0,446
10	49	4,38%	5,31%	0,435
11	15	4,34%	5,29%	0,437
Total général	**1000**	**14,79%**	**7,60%**	**0,360**

Appendix 4

Correlation matrices for the variables relative to the series Markowitz-Upcase and Markowitz-Downcase

	Markowitz. Upcase.Return	*Markowitz. Upcase.Risk*	*Number.of.funds. Upcase*	*Sharpe.ratio. Upcase*
Markowitz. Upcase.Return				
Markowitz. Upcase.Risk	0.828***			
Number.of.funds. Upcase	-0.759***	-0.861***		
Sharpe.ratio. Upcase	-0.750***	-0.889***	0.945***	

Computed correlation used pearson-method with listwise-deletion.

*Note: * = p < .05, ** = p < .01, *** = p < .001*

	Markowitz. Downcase.Return	*Markowitz. Downcase.Risk*	*Number.of.funds. Downcase*	*Sharpe.ratio. Downcase*
Markowitz. Downcase.Return				
Markowitz. Downcase.Risk	0.830***			
Number.of.funds. Downcase	-0.858***	-0.879***		
Sharpe.ratio. Downcase	-0.243***	-0.636***	0.639***	

Computed correlation used pearson-method with listwise-deletion.

*Note: * = p < .05, ** = p < .01, *** = p < .001*

Appendix 5

Univariate Statistical of the variables included the linear regression models

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Markowitz.Total.Risk	1,000	0.100	0.075	0.033	0.043	0.075	0.132	0.463
Markowitz.Total.Return	1,000	0.099	0.022	0.054	0.081	0.101	0.118	0.132
Markowitz.Upcase.Risk	1,000	0.096	0.069	0.033	0.043	0.073	0.127	0.424
Markowitz.Upcase.Return	1,000	0.098	0.021	0.055	0.081	0.101	0.116	0.130
Markowitz.Downcase.Risk	1,000	0.148	0.128	0.040	0.051	0.095	0.199	0.506
Markowitz.Downcase.Return	1,000	0.076	0.020	0.034	0.061	0.078	0.093	0.106
Sharpe.ratio.Total	1,000	0.800	0.227	0.213	0.629	0.834	1.007	1.094
Number.of.funds.Total	1,000	7.437	2.493	1	5	8	9	11
Sharpe.ratio.Upcase	1,000	0.816	0.225	0.227	0.651	0.851	1.019	1.102
Number.of.funds.Upcase	1,000	7.415	2.495	1	5	8	9	11
Sharpe.ratio.Downcase	1,000	0.360	0.139	0.007	0.229	0.404	0.488	0.524
Number.of.funds.Downcase	1,000	6.373	2.981	1	3	8	9	11

Appendix 6

Linear Interpolation Model

Linear interpolation is a method of curve fitting using linear polynomials to construct new data points within the range of a discrete set of known data po$_{in}$ts.

If two known points are given by the coordinates (x_0, y_0) and (x_1, y_1) , the linear interpolant is the straight line between these points. For a value x in the interval (x_0, x_1) the value y along the straight line is given from the equation of slopes:

$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

Solving this equation for *y*, which is the unknown value at *x*, gives:

$$y = y_0 + (x - x_0)\left(\frac{y_1 - y_0}{x_1 - x_0}\right) = \frac{y_0\,(x_1 - x) + \,y_1\,(x - x_0)}{x_1 - x_0}$$

which is the formula for linear interpolation in the interval (x_0, x_1).

Appendix 7

Cubic Spline Interpolation

We have piecewise interpolated the unknown function $r(\sigma)$ over an interval $[\sigma_0, \sigma_n]$ partitioned into subintervals $[\sigma i\text{-}1, \sigma i]$, i = 1,2, ..., n.

Cubic spline interpolation involves here replacing the points of unknown function $r(\sigma)$ with a third degree polynomial on each subinterval, so that the interpolating function is continuous, as well as its first and second derivatives over the entire interval $[\sigma_0, \sigma_n]$.

Cubic functions of interpolation f_i as thus written as follows:

$$f_i(x) = a_i\sigma^3 + b_i\sigma^2 + c_i\sigma + d_i\,,\ \ \sigma_{i-1} \le \sigma \le \sigma_i\,,\ \ i = 1,2,\ldots,n$$

We state: $f_i''(\sigma_i) = f_i''$, $i = 1,2,\ldots,n$

The continuity of the second derivative is written:

$$f_i''(\sigma_i) = f_i'' \quad et \quad f_i''(\sigma_{i-1}) = f_{i-1}''\ ,\ \ i = 1,2,\ldots,n$$

or, as the second derivative is linear:

$$f_i''(\sigma) = f_{i-1}'' \frac{\sigma - \sigma_i}{\sigma_{i-1} - \sigma_i} + f_i'' \frac{\sigma - \sigma_{i-1}}{\sigma_i - \sigma_{i-1}}$$

Therefore, by setting: hi = σi – σi-1:

$$f_i''(\sigma) = f_{i-1}'' \frac{\sigma_i - \sigma}{h_i} + f_i'' \frac{\sigma - \sigma_{i-1}}{h_i}$$

By integrating twice it comes:

$$f_i(x) = f_{i-1}'' \frac{(\sigma_i - \sigma)^3}{6h_i} + f_i'' \frac{(\sigma - \sigma_{i-1})^3}{6h_i} + r_i\sigma + s_i$$

Integration constants ri and si are determined using the known values of fi(σ) for σi and σi-1:

$f_i(\sigma_{i-1}) = y_{i-1}$, $f_i(\sigma_i) = y_i$

After some algebraic manipulations, we get:

$$f_i(\sigma) = f_{i-1}'' \frac{(\sigma_i - \sigma)^3}{6h_i} + f_i'' \frac{(\sigma - \sigma_{i-1})^3}{6h_i} + \left[\frac{\sigma_{i-1}}{h_i} - f_{i-1}'' \frac{h_i}{6}\right](\sigma_i - \sigma) + \left[\frac{y_i}{h_i} - f_i'' \frac{h_i}{6}\right](\sigma - \sigma_{i-1})$$

The functions $f_i(\sigma)$ will be fully known when we have calculated the values of f_i''.

To obtain these values, we use the conditions of continuity of the first derivatives at the interior points.

After derivation of $f_i(\sigma)$ we impose:

$$f_i'(\sigma_i) = f_{i+1}'(\sigma_i) \quad , \quad i = 1,2,...,n-1$$

to get:

$$h_i f_{i-1}'' + 2(h_i + h_{i+1}) f_i'' + h_{i+1} f_{i+1}'' = \frac{6}{h_{i+1}}(y_{i+1} - y_i) + \frac{6}{h_i}(y_{i-1} - y_i) \quad ; \quad i = 1,2,...,n-1.$$

that is a system of n-1 equations with n+1 unknown variables.

We thus have the possibility of imposing two additional conditions, obtained, for example, by the conditions at the limits of the interval [σ_0, σ_n].

If we impose the following two conditions:

$$f_1''(\sigma_0) = 0 \quad and \quad f_n''(\sigma_n) = 0$$

we get natural cubic splines.

The n-1 unknowns variables f_1'', f_2'' ,..., f_{n-1}'' are then solution of the linear system $\mathbf{A}\boldsymbol{f} = \boldsymbol{b}$, written above.

The diagonal elements of matrix A are :

$a_{i,i} = 2(h_i + h_{i+1})$, $i = 1, 2, \ldots, \text{n-1}$.

While the elements of the secondary diagonals of A are:

$a_{i+1,i} = a_{i,i+1} = h_{i+1}$, $i = 1, 2, \ldots, \text{n-1}$.

The elements of the second member b are:

$b_i = \frac{6}{h_{i+1}}(y_{i+1} - y_i) + \frac{6}{h_i}(y_{i-1} - y_i)$, $i = 1,2,\ldots,\text{n-1}$.

This system has a unique solution because its tridiagonal matrix is diagonally dominant and is therefore invertible.